SUPPLEMENTAL INFORMATION

Supplementary Methods

Isolation and Transplantation of DEN-Initiated Hepatocytes. For HCC induction, 14-day-old *Tgfbr2*$^{fl/fl}$ male littermates were injected intraperitoneally with 25mg/kg DEN (Sigma-Aldrich, St. Louis, MO). Three months after DEN injection, mouse liver was perfused via vena cava with pre-warmed (40°C) S&M buffer (0.24% Hepes, 0.019% NaOH, 0.05%KCl, 0.83%NaCl, pH7.4) and digest solution (1mg/ml Collagenase D in S&M with 0.0636% $CaCl_2$) for 5 minutes each. Then the liver was removed, washed and minced. The cell slurry was sequentially filtered through a 70μm nylon mesh. The remnant on the filter was washed off and placed in another tube containing 10mL digest solution (with 0.01% BSA) and incubated at 37°C for another 20 minutes. Then, cell suspension was filtered through another 70μm nylon mesh to remove undigested portions, followed by 2-3 times PBS washing. Finally, cells were suspended in appropriate volume of ice cold PBS for viability check and splenic injection. For splenic injection, four-week-old male C57BL/6J mice (Jackson Laboratory, Sacramento, CA) were treated with two injections of retrorsine (Sigma-Aldrich) at 50mg/kg intraperitoneally with a 2-week interval. Four weeks after 2nd retrorsine injection, each mouse received 3 × 10^6 viable hepatocytes (including TICs) in 100μl PBS via intra-splenic injection with 26-G needle. One week after transplantation, CCl4 intraperitoneal injections (0.5ml/kg, diluted in olive oil) were given to recipient mouse once a week for 3 weeks followed by one tail vein injection of GFP or Cre-expressed adenovirus (10^{11} plaque-forming unit/mouse). The recipient mice were sacrificed 12 months later to

document HCC development and tumor burden. All mice were housed under specific pathogen-free conditions with normal diet and water *ad libitum*. The protocols of mouse experiment were approved by the Tulane University Animal Care Program, following National Institutes of Health guidelines.

Culture of DEN-Initiated Hepatocytes. Initiated hepatocytes isolated from *Tgfbr2*$^{fl/fl}$ male mouse as above-described were seeded onto collagen-coated Petri dish with growth medium: 20% heat inactivated FBS, 0.01 g/L insulin, 1μM Dexamethason, 1% L-glutamine, 20 μg/L EGF and 1× antibiotic-antimycotic (ThermoFisher, Waltham, MA) in Dulbecco's Modified Eagle Medium with 4.5g/L glucose. When confluent, cells were trypsinized and passaged. Cells used in our experiments were within 8 passages in culture. For inactivation of Tgfbr2 in cultured TICs, Cre-expressed adenovirus (100 MOI) were added to the culture medium and co-incubated with cells for 72 to 96 hours. For knockdown H19 in cultured TICs, pGFP-V-RS based constructs expressing H19 specific shRNAs (shH19-1: 5'-GCAGAATGGCACATAGAAA-3'; shH19-2: 5'-GGGCAGAATGAATGAGTTTC-3') were transfected into TICs with Lipofectamin2000 (Invitrogen, Carlsbad, CA). To assess the effect of H19 on TICs tumorigenicity in mice, 3×10^6 cells were transplanted via splenic injection into the livers of male C57BL /6J mice with the same pre- and post-transplantation treatments (except adenovirus injection) as detailed in the above section.

RNA Sequencing and Bioinformatics Analysis. Total RNAs from TICs with different treatments were isolated with RNeasy Mini kit (Qiagen, Hilden, Germany) following the

manufacture's instruction. RNA samples were then processed for preparation of sequencing libraries; the libraries were sequenced on Illumina HiSeq2000 instrument generating an average of 16 million mapped reads per sample for standard RNA sequencing and 8.7 million for small RNA sequencing. RNA sequencing data were aligned to the mm9 genome assembly using Tophat (v1.4.1)-Cufflinks (v2.2.1) pipeline. The data have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE101125. For Gene Set Enrichment Analysis (GSEA), cufflinks results were used to check the genes expression pattern difference between treated and control TICs. GSEA v2.1 software from Broad Institute was used for this purpose(1). For lncRNA differential expression analysis, SAMMate (2.7.4) was used to process cufflinks results(2) with annotation file Mus_musculus.NCBIM37.66_lincRNA.gtf (http://www.ensembl.org). For transcription factor binding sites identification, H19 genes upstream sequence (-1000 to 1) was put into algorithm PROMO as a query sequence to search for potential binding sites of putative transcription factors(3). For TCGA data mining, the gene expression and clinical data of HCC patients were downloaded from TCGA database and subjected to analyses using R (version 3.4.1) and related packages.

Imaging Flow Cytometry. Primary clusters-derived cells were thoroughly fixed with 1% PFA on ice for 20 minutes, followed by staining with anti-CD44 (Cell Signaling, Danvers, MA) antibody and FITC-conjugated secondary antibody (Santa Cruz Biotechnology, Santa Cruz, CA) sequentially. Then, samples were processed by ImageStreamX® MKII Imaging Flow Cytometer (Millipore, Billerica, MA).

Chromatin Immunoprecipitation (ChIP). ChIP assay was performed using a SimpleChIP Enzymatic Chromatin IP kit (Cell Signaling Technology) according to the manufacturer's instructions. Immunoprecipitation was performed using 1 μg of anti-Sox2 antibody (Cell Signaling Technology). Rabbit IgG antibody was used as a negative control. ChIP primers were designed to amplify the different sites in H19 promoter region. The primers are as follows: mf2, 5'-AGAGCAATGGTATACGCACG-3'; f2, 5'-AAACGTGACGACTCACCCAG-3'; r2, 5'-TTATCTGGCCGCAGTGGTTT -3'; f3, 5'-ACCCATGCCCTCAAATTCCC -3'; r3, 5'-GCCCTTCATAGAACGGTGGT-3'; f4, 5'-GAAGGGCTTCAGCAGGCTAC -3'; r4, 5'-CTTAGACAGGGCCTGGCATT-3'.

Luciferase Assays. The H19 promoter fragment (-596 to -146) was cloned into the pGL4.21 vector (Promega, Madison, WI) to produce the H19 promoter-Luciferase reporter. For luciferase assays, indicted cells were transfected with H19 promoter-Luciferase vectors using Lipofectamin2000 (Invitrogen) in bio-coated 6 well plates. Renilla luciferase-expressed plasmid pRL-TK was co-transfected with H19 promoter-Luciferase vector at ratio of 1:20 as internal control. Luciferase activity assays were performed 48 hours after transfection using the Dual-Luciferase® Reporter (DLR™) Assay System (Promega).

Cell Proliferation and Spheroid Assay. For proliferation, indicated cells were cultured in serum-free medium for 24 hours to synchronize cell cycle. Then, 2×10^3 cells were seeded into each well of 96-well Matrigel-coated plates in 200μL culture medium containing 1% fetal bovine serum. Ninety microliters serum-free medium mixed with

10μL WST-1 was added to each well at indicated time points, followed by incubation for 4 h at 37°C, 5% CO_2 to develop color. After incubation, A450 nm was measured by an enzyme-linked immunosorbent assay plate reader. For spheroid formation assay, single-cell suspensions were plated at a density of 1×10^4 cells per well in 24-well Ultra-Low Attachment Plates (Corning, Tewksbury MA) and maintained in serum-free Opti-MEM™ medium (ThermoFisher) at 37°C in a 5% CO2 humidified incubator for 10 days. The numbers of spheroids were then counted and statistically analyzed.

Immunofluorescence (IF). Cells with indicated treatment or cell clusters were washed twice with cold PBS and fixed in 2% paraformaldehyde for 10 min followed by penetration with 0.1% Triton X-100 in PBS and blocked with 5% BSA at room temperature for 1 hour. Subsequently, cells were incubated with primary antibodies at 4°C overnight, followed by 3 washings and incubation with fluorescent secondary antibodies (Santa Cruz Biotechnology). After counterstaining with 4',6-diamidino-2-phenylindole (DAPI, 0.1 μg/mL) for 1 min, the coverslips were applied with mounting medium and images were captured using a fluorescence microscope. Antibodies for CD44, EpCam, Klf4, Oct4, Sox2, c-Myc, Alb, and Smad2 were purchased from Cell Signaling Technology.

Terminal Deoxynucleotidyl Transferase dUTP Nick End Labeling (TUNEL). The DeadEnd Colorimetric TUNEL System (Promega) was used for the examination of apoptotic cell death. TICs with indicated treatments were cultured in matrigel-coated plates to a confluence rate of 70%, and the cells were then processed for TUNEL

staining, according to the manufacturer's guidelines. Instead of DAB, an anti-biotin antibody (Santa Cruz) and a FITC-conjugated fluorescence secondary antibody were used to detect the apoptotic signals.

Real-Time RT-PCR. Total RNAs from cells with indicated treatment were isolated by Trizol (Invitrogen, Grand Island, NY). For first-strand complementary DNA synthesis, 1 μg total RNA was reverse-transcribed using iScript™ Reverse Transcription Supermix kit (BioRad, Hercules, CA). BioRad SsoAdvanced™ Universal SYBR® Green Supermix was used to amplify the target genes on Bio-Rad C1000 Thermal Cycler. The primers used are: for H19, 5'-AAGAGCTCGGACTGGAGACT-3' and 5'-AAGAAGGCTGGATGACTGCC-3'; for c-Myc, 5'-CGCGATCAGCTCTCCTGAAA-3' and 5'-GCTGTACGGAGTCGTAGTCG-3'; for Sox2, 5'-TTTGTCCGAGACCGAGAAGC-3' and 5'-CTCCGGGAAGCGTGTACTTA-3'; for Oct4, 5'-AGTGGGGCGGTTTTGAGTAA-3' and 5'-GGTGTACCCCAAGGTGATCC-3'; for Klf4, 5'-GACTATGCAGGCTGTGGCAA-3' and 5'-GGTAGTGCCTGGTCAGTTCA-3'; β-actin with primers 5'-CACTGTCGAGTCGCGTCC-3' and 5'-CGCAGCGATATCGTCATCCA-3' was used as a internal control.

Western Blot. Cells were lysed in RIPA buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA and 1% NP-40) supplied with phosphatase and protease inhibitors cocktails (Roche, Basel, Switzerland). After being denatured, samples were separated on SDS-PAGE gel and then transferred onto the nitrocellulose membrane (BioRad). The milk-blocked blots were incubated with different primary antibodies at 4°C overnight. After

3× washing with PBS-T, blots were incubated with IRDye 800CW or 680LT conjugated secondary antibodies (LI-COR Biosciences, Lincoln, Ne) and scanned using LI-COR Odyssey Imaging system. Primary antibodies for Tgfbr2 was purchased from Santa Cruz Biotechnology; β-actin from Sigma-Aldrich Corporation; Smads, pSmad2, c-Myc, Sox2, Akt and pAkt from Cell Signaling Technology.

Supplementary Table 1. Clinical and pathological features of 371 HCC patients

	TGF-β signature in hepatocyte			
	Early	**Late**	**Other**	**TOTAL**
Gender				
Female	21 (20.6%)	32 (42.7%)	68 (35.6%)	122 (32.9%)
Male	81 (79.4%)	43 (57.3%)	125 (64.4%)	249 (67.1%)
Age (years)				
Median (range)	62.5 (24-90)	58 (16-81)	62 (20-85)	61 (16-90)
< 60	45 (44.1%)	41 (54.7%)	85 (43.8%)	171 (46.1%)
> 60	57 (55.9%)	34 (45.3%)	109 (56.2%)	200 (53.9%)
Etiology				
HBV	20 (19.6%)	8 (10.7%)	49 (25.3%)	77 (20.8%)
HCV	9 (8.8%)	8 (10.7%)	17 (8.8%)	34 (9.2%)
HBV + HCV	2 (2%)	1 (1.3%)	1 (0.5%)	4 (1.1%)
ALD	24 (23.5%)	20 (26.7%)	34 (17.5%)	78 (21%)
ALD + Viral infection[a]	19 (18.6%)	3 (4%)	16 (8.2%)	38 (10.2%)
Other	22 (21.6%)	31 (41.3%)	68 (35.1%)	121 (32.6%)
Unknown	6 (5.9%)	4 (5.3%)	9 (4.6%)	19 (5.1%)
Fibrosis/Cirrhosis (χ^2=2.3, n.s.)				
No	24 (34.8.%)	16 (45.7%)	34 (31.8%)	74 (35.1%)
Yes	45 (63.2%)	19 (54.3%)	73 (68.2%)	137 (64.9%)
AFP[b] (χ^2=9.4, *P* < 0.01, Early vs. Other *adj.P*=0.006, Early vs. Late + Other *P*=0.003)				
< 300 ng/mL	74 (88.1%)	37 (75.5%)	102 (70.3%)	213 (76.6%)
> 300 ng/mL	10 (11.9%)	12 (24.5%)	43 (29.7%)	65 (23.4%)
Histologic Grade (χ^2=2.1, n.s.)				
G1	15 (15.2%)	10 (13.5%)	30 (15.6%)	55 (15.1%)
G2	52 (52.5%)	34 (45.9%)	91 (47.4%)	177 (48.5%)
G3	28 (28.3%)	28 (37.8%)	65 (33.9%)	121 (33.2%)
G4	4 (4%)	2 (2.7%)	6 (3.1%)	12 (3.3%)
AJCC Stage (χ^2=10.5, *P* < 0.05, Early vs. Late adj.*P*=0.017, Other vs. Late *adj.P*=0.034)				
Stage 1	55 (56.7%)	23 (33.8%)	94 (51.4%)	172 (49.4%)
Stage 2	23 (23.7%)	19 (27.9%)	44 (24%)	86 (24.7%)
Stage 3/4	19 (19.6%)	26 (38.2%)	45 (24.6%)	90 (25.9%)
Vascular Invasion (χ^2=2.8, n.s.)				
No	65 (71.4%)	35 (58.3%)	106 (65%)	207 (65.5%)
Yes	26 (28.6%)	25 (41.7%)	57 (35%)	109 (34.5%)
Recurrence (months) (log-rank test, n.s.)				
Mean ± s.d.	20.0 ± 19.9	17.9 ± 19.4	20. 5± 20.1	19.8 ± 20.1

a. Either HVB or HCV infection; b. Serum level of α-fetoprotein at procurement.

Supplementary Table 2. Clinical and pathological features of 50 HCC patients with available tumor-adjacent liver tissues

	TGF-β induced genes in HSC		
	ClusterF1	**ClusterF2**	**TOTAL**
Gender			
Female	11 (36.7%)	11 (55%)	22 (44%)
Male	19 (63.3%)	9 (45%)	28 (56%)
Age (years)			
Median (range)	66.5 (23-81)	69 (20-77)	68.5 (20-81)
< 60	12 (40%)	5 (25%)	17 (34%)
> 60	18 (60%)	15 (75%)	33 (66%)
Etiology			
HBV	6 (20%)	1 (5%)	7 (14%)
HCV	4 (13.3%)	1 (5%)	5 (10%)
ALD	4 (13.3%)	5 (25%)	9 (18%)
NAFLD	2 (6.7%)	0 (0%)	2 (4%)
Other	13 (43.3%)	11 (55%)	24 (48%)
Unknown	1 (3.3%)	2 (10%)	3 (6%)
Fibrosis/Cirrhosis (χ^2=0.004, n.s.)			
No	15 (62.5%)	7 (63.6%)	22 (62.9%)
Yes	9 (37.5%)	4 (36.4%)	13 (37.1%)
AFP[a] (χ^2=0.6, n.s.)			
< 300 ng/mL	16 (84.2%)	11 (73.3%)	27 (79.4%)
> 300 ng/mL	3 (15.8%)	4 (26.7%)	7 (20.6%)
Histologic Grade (χ^2=0.9, n.s.)			
G1	2 (7.1%)	3 (15.8%)	5 (10.6%)
G2	17 (60.7%)	10 (52.6%)	27 (57.4%)
G3	9 (30%)	6 (30%)	15 (30%)
AJCC Stage (χ^2=0.8, n.s.)			
T1	11 (36.7%)	9 (45%)	20 (40%)
T2	8 (26.7%)	6 (30%)	14 (28%)
T3/T4	11 (36.7%)	5 (25%)	16 (32%)
Vascular Invasion (χ^2=0.6, n.s.)			
No	20 (74.1%)	10 (62.5%)	30 (69.8%)
Yes	7 (25.9%)	6 (37.5%)	13 (30.2%)
Recurrence (months) (log-rank test, n.s.)			
Mean ± s.d.	17.2 ± 12.1	21.1 ± 19.0	19.2 ± 15.8

a. Serum level of α-fetoprotein at procurement.

Supplementary Table 3. Top 50 increased genes in TICs with H19 deletion

Gene symbol	Description	Fold change (shH19/shControl)
Ppp1r9a	protein phosphatase 1, regulatory (inhibitor) subunit 9A	8.00
Asb16	ankyrin repeat and SOCS box-containing 16	6.67
Ano4	anoctamin 4	6.50
Ptgis	prostaglandin I2 (prostacyclin) synthase	6.00
Mmp2	matrix metallopeptidase 2	5.50
Trim40	tripartite motif-containing 40	5.50
Apcdd1	adenomatosis polyposis coli down-regulated 1	5.00
Pabpc5	poly(A) binding protein, cytoplasmic 5	4.60
Rpl21-ps8	ribosomal protein L21, pseudogene 8	4.50
En1	engrailed 1	4.21
BC049762	cDNA sequence BC049762	4.20
Gypc	glycophorin C	4.13
Dpysl5	dihydropyrimidinase-like 5	3.83
Actn3	actinin alpha 3	3.60
Lsp1	lymphocyte specific 1	3.55
Slc5a2	solute carrier family 5 (sodium/glucose cotransporter), member 2	3.50
Ugt8a	UDP galactosyltransferase 8A	3.50
Dnah8	dynein, axonemal, heavy chain 8	3.43
Bfsp2	beaded filament structural protein 2, phakinin	3.40
Lypd1	Ly6/Plaur domain containing 1	3.40
Pcdhb21	protocadherin beta 21	3.38
Mfsd2b	major facilitator superfamily domain containing 2B	3.33
Xkr5	X-linked Kx blood group related 5	3.33
Zfp966	zinc finger protein 966	3.33
Npr1	natriuretic peptide receptor 1	3.30
S1pr1	sphingosine-1-phosphate receptor 1	3.29
Dact3	dishevelled-binding antagonist of beta-catenin 3	3.29
Abhd11os	abhydrolase domain containing 11, opposite strand	3.25
Htra3	HtrA serine peptidase 3	3.25
Npr2	natriuretic peptide receptor 2	3.24
Spn	sialophorin	3.24
Ism1	isthmin 1, angiogenesis inhibitor	3.22
Elfn1	leucine rich repeat and fibronectin type III, extracellular 1	3.21
Mfrp	membrane frizzled-related protein	3.20
Tnr	tenascin R	3.20
Gfi1	growth factor independent 1	3.17
Syt8	synaptotagmin VIII	3.04
Gad2	glutamic acid decarboxylase 2	3.00
Gper1	G protein-coupled estrogen receptor 1	3.00
Hist1h2br	histone cluster 1 H2br	3.00
Med9os	mediator complex subunit 9, opposite strand	3.00
Myocd	myocardin	3.00
Nsg1	neuron specific gene family member 1	3.00
Ramp3	receptor (calcitonin) activity modifying protein 3	3.00
Nutm1	NUT midline carcinoma, family member 1	2.93
Rapgef5	Rap guanine nucleotide exchange factor (GEF) 5	2.93
Zfp335os	zinc finger protein 335, opposite strand	2.91
Olfm2	olfactomedin 2	2.89
Nog	noggin	2.86
BC085271	protein phosphatase 2A inhibitor-2 I-2PP2A-related	2.86

Supplementary Table 4. Top 50 decreased genes in TICs with H19 deletion

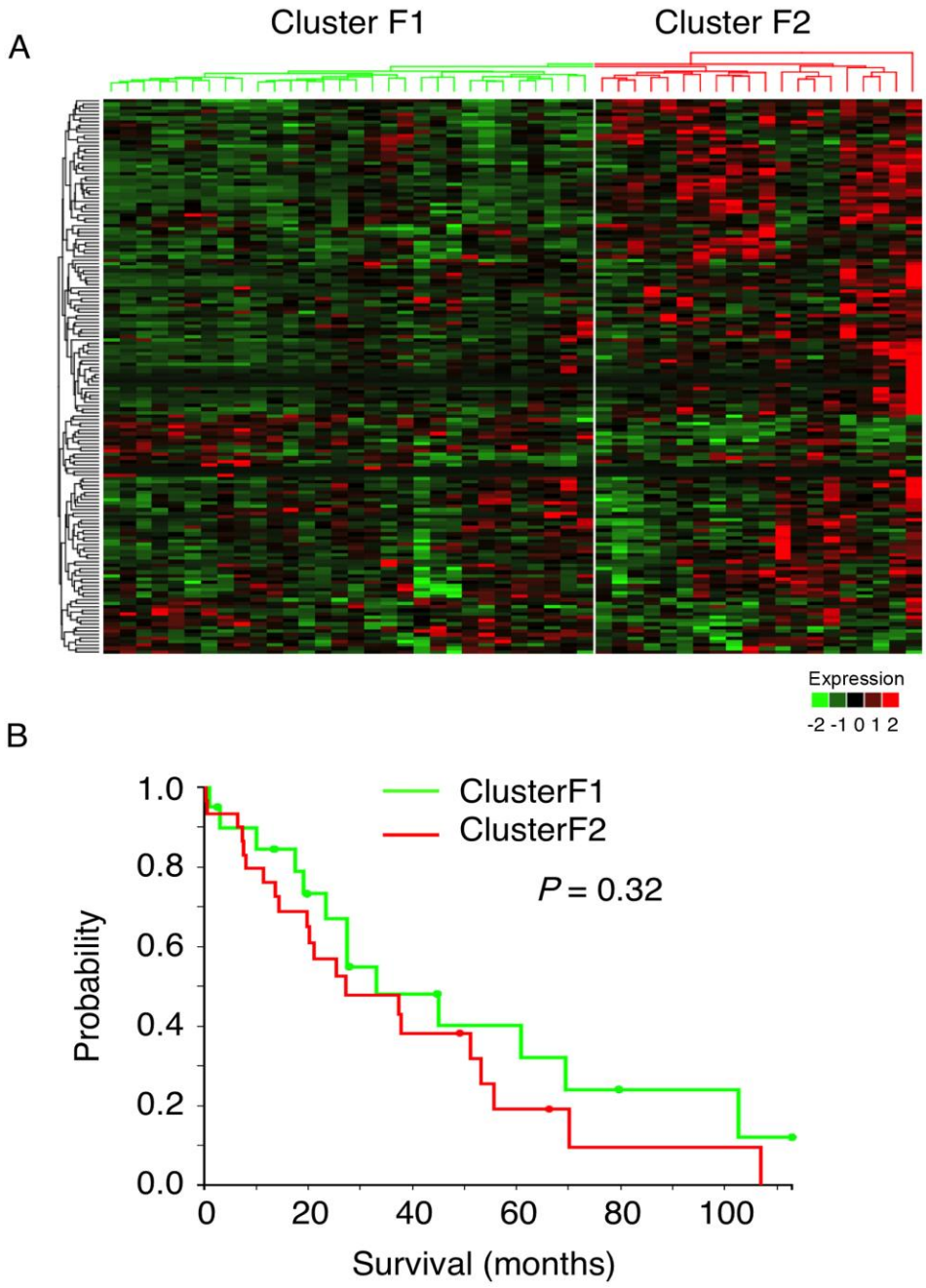
Gene symbol	Description	Fold change (shH19/shControl)
Ccdc154	coiled-coil domain containing 154	0.04
Zfp783	zinc finger protein 783	0.14
Lce6a	late cornified envelope 6A	0.17
Pramef25	PRAME family member 25	0.17
Ak8	adenylate kinase 8	0.21
Gng13	guanine nucleotide binding protein (G protein), gamma 13	0.22
Insc	inscuteable homolog (Drosophila)	0.22
Bcl2l15	BCLl2-like 15	0.24
Aplp1	amyloid beta (A4) precursor-like protein 1	0.24
Thpo	thrombopoietin	0.24
Sync	syncoilin	0.24
Rpl7a-ps9	ribosomal protein L7A, pseudogene 9	0.25
Rpl38-ps1	ribosomal protein L38, pseudogene 1	0.26
Snora19	small nucleolar RNA, H/ACA box 19	0.26
Pmepa1os	prostate transmembrane protein, androgen induced 1, opposite strand	0.26
Npas2	neuronal PAS domain protein 2	0.27
Pcp2	Purkinje cell protein 2 (L7)	0.27
Myh11	myosin, heavy polypeptide 11, smooth muscle	0.28
Tmtc2	transmembrane and tetratricopeptide repeat containing 2	0.28
Mmp16	matrix metallopeptidase 16	0.28
Ccdc106	coiled-coil domain containing 106	0.29
Fmr1os	fragile X mental retardation syndrome 1, opposite strand	0.30
Klhl10	kelch-like 10	0.30
Aoah	acyloxyacyl hydrolase	0.31
Cox6b2	cytochrome c oxidase subunit VIb polypeptide 2	0.31
Mir99b	microRNA 99b	0.31
Slc26a8	solute carrier family 26, member 8	0.31
Ccdc148	coiled-coil domain containing 148	0.32
Mir1192	microRNA 1192	0.32
Try5	trypsin 5	0.32
Apcs	serum amyloid P-component	0.32
Pglyrp3	peptidoglycan recognition protein 3	0.32
Acod1	aconitate decarboxylase 1	0.33
Igsf23	immunoglobulin superfamily, member 23	0.33
Kcnj13	potassium inwardly-rectifying channel, subfamily J, member 13	0.33
Tsnaxip1	translin-associated factor X (Tsnax) interacting protein 1	0.33
Zfp750	zinc finger protein 750	0.34
Mamstr	MEF2 activating motif and SAP domain containing transcriptional regulator	0.34
Tmem8b	transmembrane protein 8B	0.35
Ccdc146	coiled-coil domain containing 146	0.35
Dpyd	dihydropyrimidine dehydrogenase	0.35
B4galnt2	beta-1,4-N-acetyl-galactosaminyl transferase 2	0.35
Ablim2	actin-binding LIM protein 2	0.35
Fcnaos	ficolin A, opposite strand	0.35
Trim15	tripartite motif-containing 15	0.35
Tubal3	tubulin, alpha-like 3	0.36
Ccdc169	coiled-coil domain containing 169	0.36
Ccdc33	coiled-coil domain containing 33	0.37
Cxcl9	chemokine (C-X-C motif) ligand 9	0.37
Pspn	persephin	0.37

Supplementary Table 5. Top 20 increased microRNAs with H19 depletion

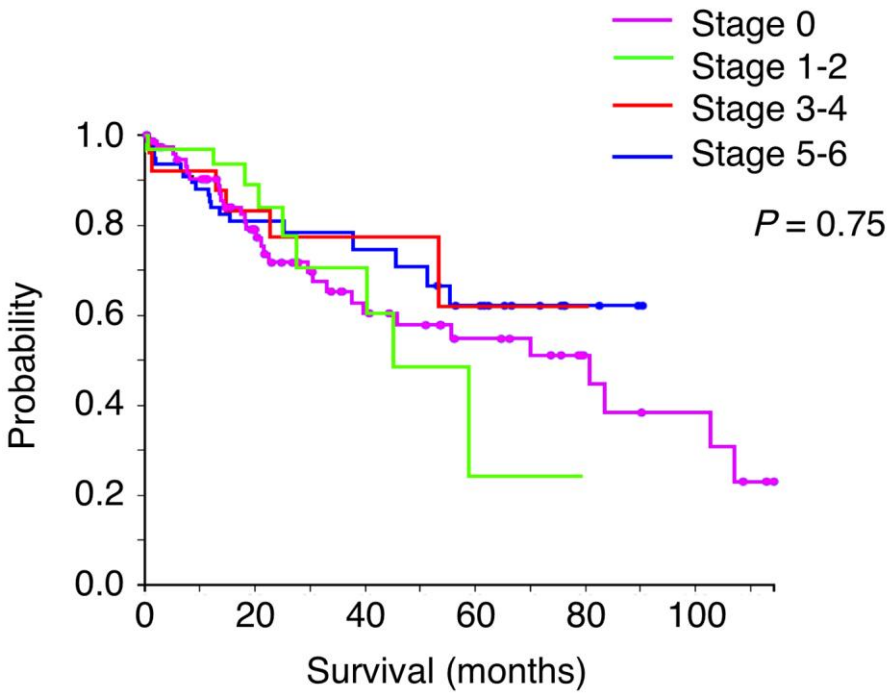
Increased MicroRNA	Fold change (shH19/shControl)
mir-205	7.71
mir-216b	6.50
mir-200b	5.50
mir-466k	5.14
mir-3112	4.00
mir-6997	4.00
mir-1960	3.63
mir-383	3.50
mir-598	3.50
mir-652	3.20
mir-200a	3.11
mir-217	2.86
mir-6916	2.80
mir-500	2.50
mir-6540	2.50
mir-466i	2.40
mir-470	2.40
mir-365-2	2.37
mir-504	2.34
mir-3098	2.31

Supplementary Table 6. Top 20 decreased microRNAs with H19 depletion

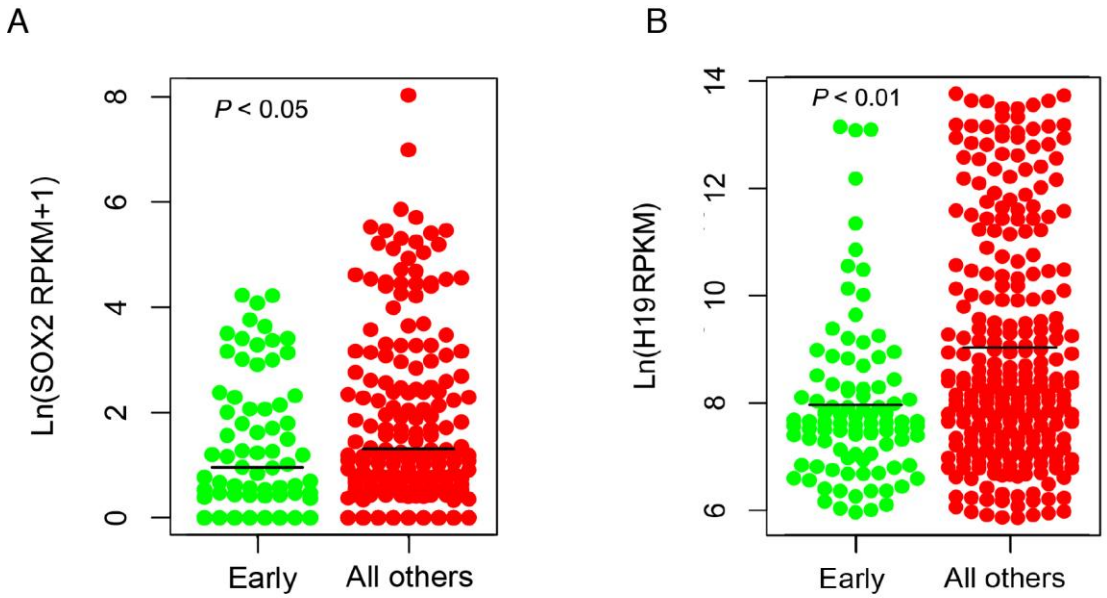
Decreased MicroRNA	Fold change (shH19/shControl)
mir-6546	0.20
mir-714	0.20
mir-592	0.21
mir-7021	0.21
mir-219c	0.25
mir-5112	0.26
mir-106a	0.30
mir-23b	0.31
mir-3470a	0.31
mir-669k	0.32
mir-7646	0.32
mir-3064	0.40
mir-6909	0.40
mir-136	0.44
mir-142b	0.44
mir-1195	0.50
mir-669m-2	0.50
mir-3470b	0.52
mir-6538	0.53
mir-702	0.54



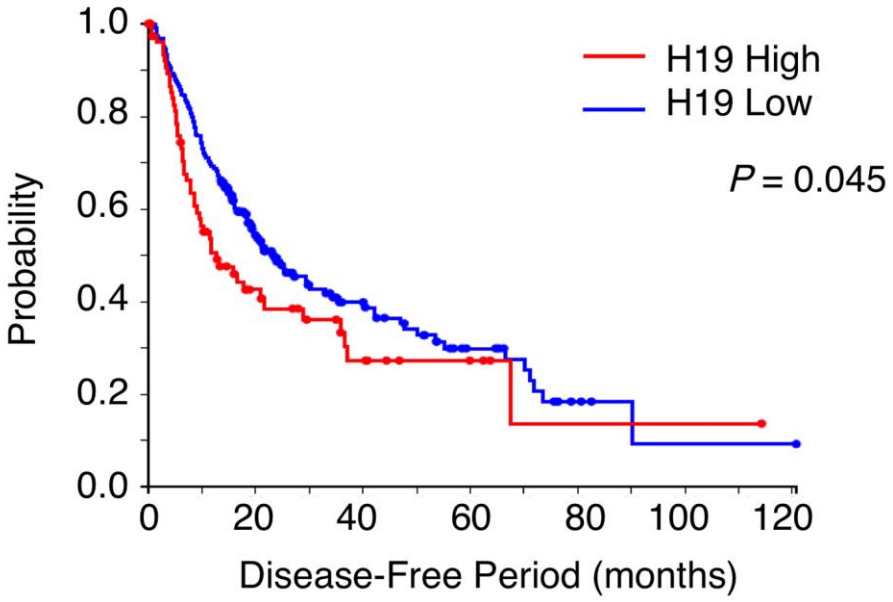
Supplementary Figure S1. (**A**) Dendrogram and heat-map overview of TGFβ-upregulated fibrogenic cell gene expression in the non-neoplastic liver tissues adjacent to HCC. Based on the expression of fibrogenic TGF-β signature genes (which include 202 up-regulated genes in TGF-β treated hepatic stellate cells) (4), the non-neoplastic liver tissues (50 cases available from the TCGA database) were divided into two subgroups: the low fibrogenic TGF-β signature (cluster F1, n=30) and high fibrogenic TGF-β signature (cluster F2, n=20). The data are presented in a matrix format in which rows and columns represent genes and patients, respectively. (**B**) Overall survival of patients with low and high fibrogenic TGF-β signatures (Kaplan-Meier and log-rank analysis).



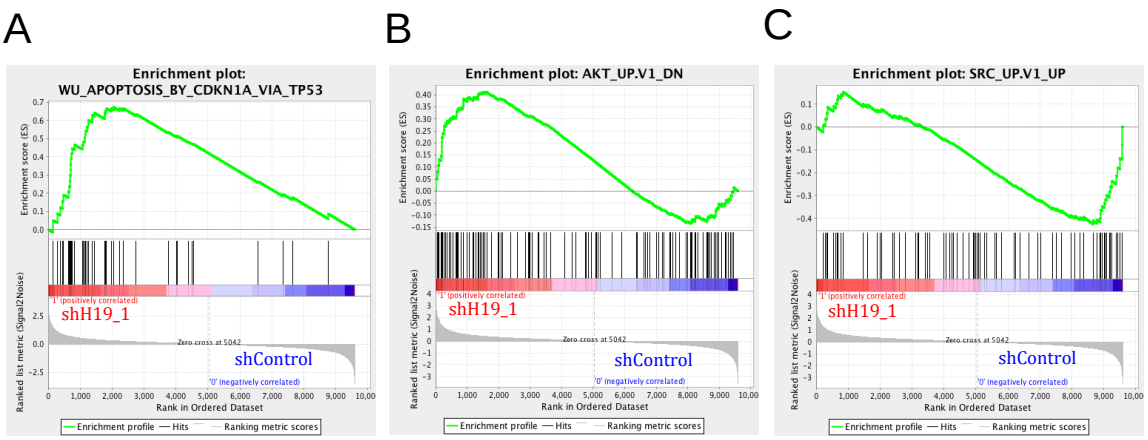
Supplementary Figure S2. Overall survival of HCC patients with different stages of fibrosis (Ishak scale) in the background livers. Kaplan-Meier plot and log-rank statistics analyses were performed in patients with different fibrosis stages available from the TCGA database. These include patients with no fibrosis (Stage 0, n=75), Stage 1-2 fibrosis (n=31), Stage 3-4 fibrosis (n=28), and Stage 5-6 fibrosis (n=79). The patients with no information on fibrosis were not included in the analysis.



Supplementary Figure S3. Dot plot analysis for Sox2 mRNA (**A**) and H19 (**B**) levels in HCC tissues. Both Sox2 mRNA and H19 levels are significantly lower in patients defined by early hepatic TGF-β signature (cluster 1, n=102) when comparing to all other patients (cluster 2 and 3, n=269). The bar indicates the mean value of each group.



Supplementary Figure S4. Disease-free survival of HCC patients with different H19 levels in tumor tissues. The TCGA HCC patients with disease-free survival data were divided into H19 high (top 25% patients based on tumor H19 level, n=79) and H19 low (bottom 75% patients, n=240) for Kaplan-Meier plot and log-rank analyses.



Supplementary Figure S5. GSEA analyses for oncogenic signatures in TICs with shRNA knockdown of H19. Knockdown of H19 led to increased expression of TP53 and CDKN1A/p21 target genes(5) (**A**), increased expression of genes which are normally suppressed by Akt(6) (**B**), and decreased expression of Src target genes(7) (**C**).

References:

1. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-15550.
2. Xu G, Deng N, Zhao Z, Judeh T, Flemington E, Zhu D. SAMMate: a GUI tool for processing short read alignments in SAM/BAM format. Source Code Biol Med 2011;6:2.
3. Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics 2002;18:333-334.
4. Tang LY, Heller M, Meng Z, Yu LR, Tang Y, Zhou M, Zhang YE. Transforming Growth Factor-beta (TGF-beta) Directly Activates the JAK1-STAT3 Axis to Induce Hepatic Fibrosis in Coordination with the SMAD Pathway. J Biol Chem 2017;292:4302-4312.
5. Wu Q, Kirschmeier P, Hockenberry T, Yang TY, Brassard DL, Wang L, McClanahan T, et al. Transcriptional regulation during p21WAF1/CIP1-induced apoptosis in human ovarian cancer cells. J Biol Chem 2002;277:36329-36337.
6. Majumder PK, Febbo PG, Bikoff R, Berger R, Xue Q, McMahon LM, Manola J, et al. mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. Nat Med 2004;10:594-601.

7. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 2006;439:353-357.